Creation of Spectral Database and ZPTW Profiles

Eric Nussbaumer [*1]

[1]*National Center for Atmospheric Research, Boulder, CO, USA*

September 2015

Abstract

This document outlines the creation of the spectral database as well as the profiles for pressure, temperature, and water vapor.

---

[*]corresponding author: ebaumer@ucar.edu +1 (303) 497-1861

1   Introduction

The spectral database and ZPTW (altitude, pressure, temperature, and water vapor) profiles are necessary pre-processing steps to retrievals. The spectral database holds information pertaining to each of the measurements. A spectral database is unique to each site

The majority of information in the spectral database comes from the OPUS file itself; however, we append meteorological data from local weather weather stations.

There are several steps in creating the spectral database:

1. Creating the initial spectral database

2. Formatting the house log data files

3. Formatting the external station weather data

4. Appending the initial spectral database with house an external station weather data

Note that not all sites have house or external station weather data.

Station	House Data	External Station Data
MLO	Yes	Yes (CMDL)
TAB	Yes	No
FL0	No	Yes (EOL)

The necessary python files are located in the sfit-processing-environment git repository.

2   Spectral Database

2.1   Initial Spectral Database

The initial spectral database file is created by running ckopus on the various raw OPUS file. A python program is created to manage the creation of the initial spectral database file (mkSpecDB.py). The program will create a new spectral database file or append an already existing file. Associated with mkSpecDB.py is an input file. The input file allows one to specify the starting and ending date to process, the station, and the various directories and files to use. In addition, one can specify additional ckopus flags to use in the ckopus call. There are logical flags which control the creating of a file which list the folders processed and whether bnr files are created. These files are located under the SpectralDatabase folder of the git repository.

Program	Description
mkSpecDB.py	Main program to create initial spectral database
specDBInputFile.py	Input file for mkSpecDB.py program

2.2   House Data

House data is data that is recorded by the FTS autonomous system, such as outside temperature, pressure, wind direction, etc. The format of this data has changed for each station over time as the instrument gets modified or upgraded. A python program (station_house_reader.py) is created to read the various formats and create a standardized file. There is one file for each year. There are no input files for the station_house_reader.py program. The time range, station identifier, and directories are specified directly in the program under the main function. An excel spreadsheet describes the various formats for the house log files for MLO and TAB.

Program	Description
station_house_reader.py	Main program to read house data files
HouseReaderC.py	Supporting program with formats of previous house data files
HouseDataLog.xlsx	Excel file with format of house log files

These programs are located in the ExternalData folder of the git repository.

2.3   External Station Data

There are currently two external station data sources used (EOL for FL0, and CMDL for MLO) only the EOL data needs to be pre-processed. The original format of this data is in netcdf files. The program read_FL0_EOL_data.py reads the daily netcdf files and creates a yearly text file. There are no input files for read_FL0_EOL_data.py program. The year of interest and directories of data are specified directly in the program under the main function. The program pullAncillaryData.py pulls the CMDL and EOL data from each individual ftp site.

Program	Description
pullAncillaryData.py	Program to automatically pull EOL and CMDL data
read_FL0_EOL_data.py	Main program to read EOL and CMDL data

These programs are located in the ExternalData folder of the git repository.

2.4   Append Spectral Database File

The final step is appending the initial spectral database file with the house and external station weather data. A python program was created to accomplish this (appendSpecDB.py). The program appendSpecDB.py reads in the initial spectral database file. It then searches the house and external station files for weather data at the time of observation, plus a certain number of minutes specified by the user. The mean of the data collected is calculated and a new spectral database file is created. If no data is present missing values are used. Associated with appendSpecDB.py is an input file. The input file allows one to specify directories and files, year to process, station, how many minutes to use for averaging, and whether to create a comma separated or pre-specified formatted new spectral database file.

Program	Description
appendSpecDB.py	Program to create the append spectral database file
appndSpecDBInputFile.py	Input file for appendSpecDB.py

Note:A warning message will often appear when running this program originating from the python numpy module. This warning is a result of numpy taking the mean of an empty array. This is handled by the main program.

2.5   Steps for Creating Spectral Database

1. Download external station data using pullAncillaryData.py

2. Create initial database file by running mkSpecDB.py

3. Format house data by running station_house_reader.py

4. Format external station data by running read_FL0_EOL_data.py

5. Create appended spectral database file by running appendSpecDB.py

3   ZPTW Profiles

The pressure, temperature, and water vapor profiles can be created from several outside sources. Temperature, pressure, and water profiles are taken from NCEP nmc data; while currently only water profiles are taken from ERA-Interim data. Both NCEP and ERA-Interim data are interpolated with WACCM data to reach 120km vertical height. The profiles are daily averages and they reside in the raw data directories (directories with bnr files).

3.1   WACCM & NCEP Profiles

Profiles from WACCM are monthly means. These monthly means have been pre-calculated from the WACCM model. The profiles from NCEP come from the NCEP nmc data which are daily averages. The program MergPrf.py creates either WACCM and NCEP altitude, pressure, and temperature profiles as files named ZPT.nmc.120. Water files are also created (w-120.v1) using the NCEP data. When NCEP data is not available the full WACCM profiles are used. These files are stored in the raw data directories.

Program	Description
MergPrf.py	Main program to create ZPT and water files from NCEP data

3.2   ERA Interim Water Profiles

The ERA-Interim daily profiles are calculated from 6 hourly data. Both the 6 hourly and daily data for profiles are created. The ERA-Interim data is housed locally at NCAR in the CISL Research Data Archive. Information regarding the data set can be found at: ds627.0.

There is a three month lag between the current date and when the data becomes available. The data is hosted on /glade/ and can be accessed through the data-access.ucar.edu server. The data can be found at: /glade/p/rda/data/ds627.0/ei.oper.an.pl/. The following steps should be used to pre-process the data:

1. Copy over the data from glade

2. Convert GRIB format files to NetCDF files using cnvrtNC.py

3. Create water profiles using ERAwaterPrf.py

Program	Description
cnvrtNC.py	Program to convert ERA-Interim GRIB files to NetCDF files
ERAwaterPrf.py	Program to extract daily averaged water profiles from ERA-Interim

4   Retrieved Water Profiles

For all sites (MLO,TAB, and FL0) water is retrieved when available. This water can be used as a prior for other retrievals. The program retWaterPrf.py creates w-120.YYYYMMDD.HHMMSS.v99 for each retrieval. These files are stored in the raw data directories (/data1/tab,mlo,fl0/). A daily average of these profiles can be created using the program retWaterPrfDaily.py. These daily averages are also stored in the main data directories.

Program	Description
retWaterPrf.py	Program to create water profiles from water retrieval
retWaterPrfDaily.py	Program to create daily average profiles from water retrievals